COMMITTEE FOR PROPRIETARY MEDICINAL PRODUCTS (CPMP)

POSITION PAPER ON TERMINOLOGY IN PHARMACOGENETICS

DISCUSSION AT THE PHARMACOGENETICS EXPERT GROUP	April/September 2001
SUBMISSION TO CPMP	November 2001
RELEASE FOR CONSULTATION	December 2001
DEADLINE FOR COMMENTS	June 2002
DISCUSSION AT THE PHARMACOGENETICS EXPERT GROUP	September 2002
ADOPTION BY THE CPMP	November 2002
DATE FOR COMING INTO OPERATION	June 2003

1.   Introduction

Pharmacogenetic research started from the observations that not all subjects respond in the same way to the same medicine and that these differences between individuals may be caused partially by their genetic profile.

Today the drug development programmes consider (usually for practical reasons) the subjects as coming from a rather homogenous population since it is not possible to accommodate fully in the drug development programme the whole range of inter-individual variability observed within a population. When differences in drug response are anticipated, e.g. in subjects with renal or hepatic disease, or with age-related differences, then studies are requested in the specific subgroup identified.
The contribution of genetic influences to variability in drug response often far exceeds that of any other variable and is what the science of pharmacogenetics aims to unravel. The analysis of a broad set of genetic variations may show that a genotypically defined subgroup of subjects may have a higher probability of responding to a certain drug differently from others in the population. The overall genetic profile may vary according to ethnicity.

As a result of the development within the areas of genetics and genomics, changes are likely to occur in the way drug development is currently being conducted and the way medicines will be used.

The use of terms that are harmonised and widely accepted by the stakeholders would contribute greatly to clarity in the dialogue. At present there is not an agreed set of working definitions crucial for pharmacogenetic clinical research. This is urgently required for protocols and guidelines addressing pharmacogenetic testing to ease communication particularly between ethics committees, investigators and subjects.

Following extensive consultation, the CPMP has agreed on a specific set of definitions directly relevant to the current practices in clinical research, with the understanding that they may have to be revisited in the light of future scientific advance and taking into account emerging legislation. The definitions discussed hereafter are highly relevant to the scenario of individual clinical protocols including pharmacogenetic testing; the principles might however be relevant also for trials involving testing other than pharmacogenetics.

The terms "pharmacogenetics" and "pharmacogenomics" as well as the terms used in the handling of samples and data for pharmacogenetic testing have been defined from the scientific-technical point of view.
The same definitions, following appropriate consultation will then be written in lay-terms and made available in all EU official languages to constitute a useful technical asset for regulatory authorities, ethics committees, health professionals and subjects when confronted with pharmacogenetic testing protocols and consent documents for medicinal product clinical trials.

2.   Scope

This position paper focuses on a specific set of critical terms that are frequently used in protocols for pharmacogenetic testing and that are relevant to define appropriate levels of protection for the privacy of the subjects when describing how the results and samples will be used in clinical trials.

The choice of the level depends on the extent to which it is desired or considered possible to link the data and samples to an identifiable subject and corresponds to the defined category of sample linkage.

The most appropriate level for a particular study depends on the nature of the research, the intended use of the data, the regulatory and legal environment and the specific concerns of the investigator and study sponsor. This choice must respect the needs for the privacy of subjects participating in a clinical study.

Generally, the greater the subject privacy in a study, the less are the opportunities for the subject after sample collection and pharmacogenetic testing have been performed to withdraw the individual samples from further analyses or to receive individual results from the study.

Privacy of information, control over the use of samples, and knowledge of study results may all contribute to a subject's willingness to take part in a study, and as a consequence the choice of process may significantly affect enrolment in a clinical trial in which pharmacogenetic testing is planned.

Sample coding procedures should be documented according to Good Clinical Practices (GCPs) and as provided for by relevant EU directives and accompanying guidance documents. Primary study data and original study-related records should be accessible to the competent regulatory authority in order to validate the evidence that is reported. While the regulatory authority can accept different levels of documentation, depending on the particulars of the study and the availability of other evidence or records, there may be times when it is necessary to link a clinical outcome to a particular patient. In principle, there is a framework for protecting patients enrolled in clinical trials now, and this framework may be adequate, perhaps with small changes, to apply to clinical pharmacogenetic trials.

Complete anonymity of the subject without any possibility of linking the samples/data to an individual will have great impact on the usefulness of the results and on what aspects might be verified during a GCP inspection from a competent authority or a sponsor audit. The individual subject record is an important component of data for submission to regulatory agencies and so the use of data from a study involving anonymised samples might not be acceptable for the submission of a claim to be included in the label of a drug or clinical diagnostic assay.

In designing clinical trials, investigators and sponsors should attempt, in consultation with competent authorities and ethics committees, to find the optimum balance between achieving the aims of the study and protecting the subject's safety or right to privacy.

It is recognised that DNA data unique to a subject could potentially be used to reconstruct a link between a subject's medical record and genotype information. Procedures should ensure that in order to respect the subject's wishes and privacy, such links are not reconstructed. For the same reasons, it is further recommended that the code should comprise randomly assigned numbers/letters and should not be based on protocol and site number (and perhaps gender) because if a particular site has included only a few subjects, it might be theoretically possible to reconstruct a link to individual subjects.

3. Pharmacogenetics and Pharmacogenomics

There is at present no consensus in the literature on the definitions of "pharmacogenetics" and "pharmacogenomics". Actually the terms are frequently used interchangeably. The achievement of widely accepted working definitions of the two would be a useful first approach to applying pharmacogenetics and pharmacogenomics in clinical trials. It is important to single out pharmacogenetics and pharmacogenomics from the wider field of genetic testing as the latter encompasses different level of concerns especially in terms of sensitivity of sample handling, data and trial results management.

Pharmacogenetics is the study of interindividual variations in DNA sequence related to drug response.

Pharmacogenomics is the study of the variability of the expression of individual genes relevant to disease susceptibility as well as drug response at cellular, tissue, individual or population level. The term is broadly applicable to drug design**,** discovery, and clinical development

4. Definitions applicable to DNA samples and data in clinical trials including pharmacogenetic testing

Different terminologies relate to the collection of human samples for pharmacogenetic research and the management of the data therefrom. The set of terms described in this paper are a key to correct handling of the samples and the data and to transparency of communication among industry, ethics committees, regulatory authorities and subjects about the pharmacogenetic approach in clinical research, regulatory assessment of medicinal products and clinical practice.

The processes by which samples and data are collected, labelled and stored have a direct effect on how the samples and the results obtained can be used in the future and on the obligations of the investigator and sponsor to the sample subject. This pertains particularly to situations when a subject withdraws his or her consent to further participation in a study and affects the possibility to return information to the subject or his/her physician, the possibility to withdraw a sample from future analyses and verification of data ascribed to a subject in reports and regulatory submissions. Additionally, the readiness and willingness with which a subject would or would not want to take part in a study may be affected by such factors as the uses of the results, the nature of the information the subject might receive, and the perceived risk resulting from disclosure of genetic information to third parties.

Five definitions (See table 1) for the labeling and coding of pharmacogenetic samples and data are proposed describing direct implications for the handling methodology of samples for pharmacogenetic testing and corresponding consequences for the level of privacy protection and use of the information for regulatory purposes. Duration of retention of the sample or its destruction needs to be defined in the protocol and in the consent form. Otherwise, if and when relevant, the timepoint and the procedure for anonymisation of the sample itself should be defined in these documents.

4.1 Identified samples and data
are those labeled with personal identifiers such as Name or Social Security Number.

Identified samples and data are treated in much the same way as those acquired in everyday medical practice. Because the sample and the data generated from it are directly traced to the subject, it is easy to withdraw the sample or the data from the study, update subject information, and return results to the subject. Also, at an inspection of the study it will be possible to verify the connection between the subject and the reported results. On the other hand, since a subject's genotyping results are directly linked to the subject's identity, the use of identified samples offers no extra privacy protection in addition to those generally provided.

Identified samples and relevant data might be coded at the given point in time in order to provide for extra long-term privacy protection following the closure of the trial.
The protocol should also specify when and whether the samples and data might be destroyed or anonymised.

4.2 Single coded samples and data
are those to which a single specific code is attributed for protecting individuals. It is recommended that the code should comprise randomly assigned numbers/letters

The investigator stores the key connecting the code of the sample to the individual's data. This step separates the subject's identity from the results of the pharmacogenetic analysis. The researcher with knowledge of the pharmacogenetic data would not have ready access to the identity of the subject.
Only breaking the code can reveal the subject's identity.
It is possible to withdraw a subject's sample for prospective use or return individual results to the subject or physician if desired.
The maintenance of a link between the subject and the pharmacogenetic information by a single code allows verification of data ascribed to an individual subject. Because the investigator who has coded the sample might also have access to the pharmacogenetic data, the safeguards of the subject's privacy, including doctor-subject confidentiality, are equivalent to those in current clinical trials practice.

4.3 Double-coded samples and data
have an additional privacy safeguard imposed by the use of a second coding system. Adding an additional code to the samples and data provides further protection.

The investigator who only knows the first code does not know this second code. In this way, anyone with knowledge of the pharmacogenetic results can only trace a subject identity to a coded identifier but no further, unless a key is used to link the codes between the data set with subject identifiers and the data set containing the pharmacogenetic information.
The code key linking the double coded pharmacogenetic samples and information is kept by a third party. This should not be the investigator in possession of the key linking coded sample and/or information to the subject.
The key to the double code might be maintained by the sponsoring organisation, in areas entrusted with maintaining confidential information (e.g. legal, quality assurance, clinical statistics) under strict operating procedures. Alternatively, the key might be held by an external entity, such as governmental agency, legal counsel, or other qualified third party not involved with the research.

The individual can only be linked with the sample or data obtained from it by bringing the two code keys together. Although the samples do not carry any information on the identity of

the subject, it is still considered to be possible to identify the subject as long as both code keys exist.

As with single coded samples, the existence of a link between the pharmacogenetic data and the subject's identity makes it possible to withdraw a sample or data (up to the time the results stemming from that data are reported), update subject information, return results and inspect the process. However, the conditions under which the pharmacogenetic information might be linked back to the subject's identity for any purpose are determined strictly by the specifics of the research protocol.
These conditions should be explicitly described in each protocol, and included within the subject's informed consent.

4.4 Anonymised samples and data

are for practical purposes <u>double coded</u> samples where the key linking the first and/or second code is deleted. They may be also previously single coded samples where the single code key is destroyed or even previously identifiable samples where the name/identifier is removed.

Anonymised samples and data do not carry any longer personal identifiers. Once the linking key has been deleted, information related to the subject's identity is no longer linked to data related to the pharmacogenetic results. This offers an additional level of security to the individual's data.

After anonymisation it is not possible to withdraw a subject's sample from analyses, to update subject information for further use, or to return any individual results to the subject or the subject's physician. Similarly, it also is not possible to inspect the study to determine that pharmacogenetic data is accurately correlated to a specific subject.

There will be times when stored samples may provide a regulatory agency additional information related to clinical outcome. The ability to link individual data to a patient will be essential in some circumstances and anonymised samples would be a problem.

In general, anonymised samples are well suited to research studies in which hypotheses are generated, but may be less so for clinical trials on which label claims are based.

4.5 Anonymous samples and data

are those that do not have any link whatsoever between the sample and the individual identity.

Anonymous samples may have population information (e.g., the samples may come from subjects with diabetes) but no individual data that might allow the identity of the subject to be traced. The clinical information is limited to broad categories of data, such as "male, age 50-55, cholesterol > 240 mg / dl". In many instances, the sample has no clinical data at all.
This situation is applicable in cases where the population is large enough and measures are taken in building up the code (see recommendations on page 3 on reconstructing a link).
Anonymous samples are useful in some types of pharmacogenetic studies.

Table 1. Summary table of the five terms of sample labelling.
:

Sample Labelling Category	Link Between Subject Identity and Pharmacogenetic Data	Records Identifiable for Clinical Monitoring	Actions Possible if subject withdraws Consent	Return of Individual Results to Subject	Scope of Subject Privacy protection
Identified	Yes, directly	Yes	Sample can be withdrawn with immediate effect for any prospective use	Possible	Similar to general healthcare confidentiality
Single coded	Indirectly, via code key	Yes, via protocol-specified procedures	Sample can be withdrawn with immediate effect for any prospective use	Possible	Standard for clinical research Conforms to principles of GCP
Double-coded	Very indirectly, via two sets of code keys	Yes, via protocol-specified procedures	Sample can be withdrawn with immediate effect for any prospective use	Possible	Double code offers added privacy protection over single code
Anonymised	No. Key(s) identifying the link between pharmacogenetic data and the identity of the subject is deleted	No	Sample and data are not identifiable. Sample cannot be withdrawn once key is deleted	Not possible	Pharmacogenetic data not linked to individuals
Anonymous	No	No	None	Not possible	Complete